

ABSTRACT OF THE DISCLOSURE

To provide a document extracting device, a document extracting program, and a document extracting method, having a low cost and a small amount of computation required to extract documents, a document extracting device includes similarity computing device computing all degrees of similarity between a plurality of documents to be candidates for extraction, and document extracting device to extract a combination of documents whose sum of the degrees of similarity between the documents computed by the similarity computing device is the smallest when any number of documents are extracted from among a group of the documents. As a result, since the cost required for works of giving keywords to the respective documents is not required to extract the documents, and even when the number of documents is increased, the amount of computation required to extract the documents is not increased extremely.